(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 November 2001 (08.11.2001)

PCT

(10) International Publication Number
WO 01/83806 A1

(51) International Patent Classification[7]: C12P 21/04

(21) International Application Number: PCT/US01/14349

(22) International Filing Date: 2 May 2001 (02.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/563,222 2 May 2000 (02.05.2000) US

(71) Applicant *(for all designated States except US)*: **EPICYTE PHARMACEUTICALS, INC.** [US/US]; Suite 150, 5810 Nancy Ridge Drive, San Diego, CA 92121 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HIATT, Andrew, C.** [US/US]; 660 Torrance Street, San Diego, CA 92103 (US). **HEIN, Mich, B.** [US/US]; 1355 Santa Margarita Drive, Fallbrook, CA 92028 (US).

(74) Agents: **CHRISTIANSEN, William, T.** et al.; Seed Intellectual Property Law Group PLLC, Suite 6300, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/83806 A1

(54) Title: IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS

(57) Abstract: Arrays of immunoglobulin binding proteins in plants or plant cells are provided. Such arrays comprise plants or plant cells transformed with polynucleotides encoding multiple different immunoglobulin binding proteins, or polypeptide components thereof. Methods are further provided for genetic segregation of the transformation events such that each transformant in an array is capable of producing progeny capable of expressing one or more immunoglobulin binding proteins, including multi-subunit proteins.

pairs. This combinatorial aspect of random libraries makes expression of these libraries in other organisms unfeasible. Newer technologies involve transgenic mice expressing antibodies from human chromosomal segments which can be used to generate hybridoma arrays expressing human antibodies.

Arrays formed in B-lymphocytes, phage infected bacteria or transgenic animals have been useful within certain immunoglobulin molecule screens, but difficulties have been encountered with producing large quantities of immunoglobulin molecules in these cells. Large-scale production of immunoglobulin molecules from any of the traditional organisms is typically very expensive. Further, phage infected bacteria are incapable of providing the variety of immunoglobulin molecule structures that may be desired. Similarly, the usefulness of transgenic animal cells has been limited by the susceptibility of such cells to infection with viruses or other microorganisms.

For economic and other reasons, it would be desirable to use genetically engineered plants as the primary vehicle for the discovery of immunoglobulin molecules, as well as for the ultimate production of immunoglobulin molecules to be used in industrial, clinical or research applications. The advantages of plants for production of immunoglobulin molecules include a low cost of production, relatively low capital investment compared to fermentation systems, the absence of animal viruses and prions, production of the immunoglobulin molecule in a biochemical background of defined proteins such as seed proteins, ease of storage and transport, and a facile scale-up to unlimited quantities of raw material. It would also be desirable to be able to express a library of binding proteins that is not derived from a combinatorial process of randomly paired heavy and light chains.

It is known that immunoglobulin molecules can be expressed in a variety of eukaryotic hosts including plant cells. A wide variety of structural genes have been isolated from mammalian cells and viruses, joined to transcriptional and translational initiation and termination regulatory signals from a source other than the structural gene, and introduced into plant hosts in which these regulatory signals are functional. Among those host cells that have been transformed with individual immunoglobulin

Accordingly, there remains a need in the art for methods for generating arrays of immunoglobulin molecules in plants and plant cells, as well as other eukaryotic organisms and cells. The present invention fulfills these needs and further provides other related advantages.

Summary of the Invention

Briefly stated, the present invention provides methods for the production of arrays of biologically or physiologically active immunoglobulin binding proteins in eukaryotic cells. Within certain aspects, methods are provided for preparing an immunoglobulin binding protein array in plant cells, comprising the steps of: (a) transforming a population of plant cells with a library of at least two different polynucleotides encoding different immunoglobulin binding protein (IgBP) polypeptides that: (i) specifically bind to a ligand with a $K_D<10^{-6}$ moles/liter; or (ii) form one or more disulfide bonds with one or more polypeptides in the transfected cell, to generate a binding protein that specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter; wherein the IgBP polypeptides (i) comprise four framework regions (*e.g.*, human or murine) alternating with three complementarity determining regions and (ii) comprise at least one peptide sequence having at least 75%, preferably at least 95%, sequence identity to a framework region of a native IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda immunoglobulin molecule; and wherein the IgBP polypeptides are not detectably expressed by the plant cells prior to transformation; and (b) selecting transformed plant cells, and therefrom preparing an IgBP array in plant cells. Each IgBP polypeptide may be a functional IgBP or an IgBP component (*e.g.*, a portion of an immunoglobulin molecule selected from the group consisting of heavy chains and fragments thereof, light chains and fragments thereof, J chains and secretory components) that, upon disulfide linkage to one or more IgBP components encoded by other polynucleotides in the library, forms a functional IgBP. Within certain specific embodiments, a library employed in such methods commprises at least 10, 100, 1,000 or 10,000 different polynucleotides.

Resulting C_HBPs may be assembled, for example, from four alpha chains and one J chain, from twelve mu chains and/or from ten mu chain and at least one J chain. C_HBPs or components thereof may, but need not, further comprise one or more portions of immunoglobulin molecules selected from the group consisting of J chains, secretory components and light chain constant regions. The C_HBPs may accumulate in an intracellular compartment of the cells or may be secreted from the cells.

Within further aspects, methods are provided for preparing a heavy chain binding protein array in eukaryotic cells, comprising the steps of: (a) exposing multiple copies of a polynucleotide encoding a native heavy chain to a mutagen, such that random or site-directed mutagenesis of the polynucleotide occurs, resulting in a library of heavy chain variants; (b) transforming a population of eukaryotic cells with the library of heavy chain variants; and (c) growing the transformed cells on a medium that permits assembly of C_HBPs, wherein each C_HBP comprises at least four combining sites; and therefrom preparing a C_HBP array in eukaryotic cells.

Methods are further provided for preparing a plant C_HBP array, comprising the steps of: (a) transforming a population of plant cells with a library of at least two different polynucleotides, wherein each polynucleotide encodes a different C_HBP component that forms one or more disulfide bonds with one or more polypeptides in the transformed cell to generate a C_HBP that specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter, wherein each component: (i) comprises an amino acid sequence that is at least 75%, preferably at least 95%, identical to a constant region tailpiece of a mu or alpha chain of a native immunoglobulin heavy chain; and (ii) comprises multiple combining sites, wherein all of the combining sites satisfy the same one of the following requirements: (1) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin light chain variable region or (2) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin heavy chain variable region; (b) growing the transformed plant cells on a growth medium to form plants; and (c) sexually crossing the plants to generate progeny, such that the progeny comprise polynucleotides encoding C_HBP components sufficient to form a functional C_HBP that comprises at least four combining sites; and therefrom preparing a plant

or alpha chain of a native immunoglobulin heavy chain; (b) comprises at least four combining sites, wherein all of the combining sites satisfy the same one of the following requirements: (i) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin light chain variable region; or (ii) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin heavy chain variable region; and (c) either (i) specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter; or (ii) forms one or more covalent bonds with one or more polypeptides in a cell, to generate a C_HBP that specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter; and (d) differs in amino acid sequence from other C_HBPs in the array.

Methods are further provided for preparing a heavy chain binding protein array in eukaryotic cells, comprising the steps of: (a) exposing multiple copies of a polynucleotide encoding a native heavy chain to a mutagen, such that random or site-directed mutagenesis of the polynucleotide occurs, resulting in a library of heavy chain variants; (b) transforming a population of eukaryotic cells with the library of heavy chain variants; and (c) growing the transformed cells on a medium that permits assembly of C_HBPs, wherein each C_HBP comprises at least four combining sites; and therefrom preparing a C_HBP array in eukaryotic cells.

Within further aspects, the present invention provides C_HBPs that: (a) comprise an amino acid sequence that is at least 75% identical to a constant region tailpiece of a mu or alpha chain of a native immunoglobulin heavy chain; (b) comprise at least four combining sites, wherein all of the combining sites satisfy the same one of the following requirements: (i) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin light chain variable region; or (ii) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin heavy chain variable region; and (c) either (i) specifically bind to a ligand with a $K_D<10^{-6}$ moles/liter; or (ii) form one or more covalent bonds with one or more polypeptides in a cell, to generate a C_HBP that specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All

organisms or clones of cells can be identified rapidly, enabling the easy access to an economical, high yield process for the large scale production of a desired IgBP.

GLOSSARY

Prior to setting forth the present invention in detail, definitions of certain terms used herein are provided.

Immunoglobulin binding protein (IgBP): An immunoglobulin binding protein (i) comprises an amino acid sequence that is at least 75% identical to at least one framework region of a native immunoglobulin molecule (*e.g.*, IgM, IgG, IgA, IgD, IgE, IgY kappa or lambda) and (ii) is a functional binding protein. Framework regions are described below, under "Immunoglobulins." A protein P is a functional binding protein if (1) for one molecular, ionic or atomic ligand A the K_D (P, A)<10^{-6} moles/liter (preferably <10^{-7} moles/liter), where K_D (X,Y) = [X][Y]/[X:Y], and (2) for a different molecular, ionic or atomic species B, K_D (P, B)>10^{-4} moles/liter. Such a protein P is said to specifically bind A. Immunoglobulin binding proteins (IgBPs) generally function as a binding protein by virtue of the properties of a sequence of amino acids comprising a combining site, as defined below. An IgBP may comprise a single immunoglobulin chain or fragment thereof, multiple identical immunoglobulin chains or fragments thereof, or multiple non-identical immunoglobulin chains or fragments thereof. IgBPs include, for example, single chain antigen binding proteins, Fabs and Fvs. Also included are heavy chain binding proteins (C_HBPs), discussed in greater detail below.

Component of an IgBP: a polypeptide capable of forming one or more covalent bonds (preferably disulfide bonds) with one or more other polypeptides to generate a functional binding protein. A component is not itself a functional binding protein. For example, a multimeric antibody is considered an IgBP, and the polypeptide chains that are joined by covalent bonds to form an antigen binding site are considered to be IgBP components. Examples of such components include but are not limited to heavy chains and fragments thereof, light chains and fragments thereof, J chain and fragments thereof, and secretory component and fragments thereof.

known immunoglobulin sequence over at least 50 consecutive amino acid residues of a constant region, if the immunoglobulin contains a constant region, and 75% identical to at least 25 consecutive amino acids of a variable region. Immunoglobulins may comprise multiple immunoglobulin components. Examples of such components include, but are not limited to, heavy chains and fragments thereof, light chains and fragments thereof, J chain and fragments thereof, and secretory component and fragments thereof. An immunoglobulin is generally identified by its binding specificity for a unique epitope.

Immunoglobulins are composed of the linear combination of a basic domain structure. Each domain contains two beta-pleated sheets, one beta sheet consisting of four beta strands, the other consisting of three beta strands. The two beta sheets are covalently linked by a disulfide bond. Antibody variable regions contain three sequences termed complementarity determining regions (CDR) within which are amino acid sequences of high variability when comparing numerous variable region sequences. Flanking each CDR are sequences of lesser variability termed framework regions (FR), of which there are four. The positions of the CDRs primarily coincide with the loops between beta strands, and conversely the FRs correlate with the beta strands themselves of the basic domain structure. For example, CDR1 (closest to the amino terminus of the immunoglobulin polypeptide) lies between beta strands 4-2 and 3-1, and CDR2 is between 3-1 and 4-4. The three stranded beta sheets of variable regions are the contact areas between the light chain and heavy chain variable regions. The following table illustrates the typical structure of variable regions.

Table 1

Amino Acid Residues Associated with Framework Regions and Complementarity Determining Regions of Immunoglobulin Light and Heavy Chain Variable Domains

Segment	Light chain amino acids	Heavy chain amino acids
FR1	1-23	1-30
CDR1	24-34	31-35
FR2	35-49	36-49

LEU/ILE/VAL/PHE/THR/ALA-THR/SER/ILE/ALA; (6) CDR 3 (heavy chain): start is always 33 residues after end of CDR 2 (always 2 after a CYS); residues before always CYS-XXX-XXX (typically CYS-ALA-ARG); residues after always TRP-GLY-XXX-GLY (SEQ ID NO: 4).

5

Table 2

CDR consensus sequences

Light chain sequences.

Type	CDR 1	CDR 2	CDR 3
Human kappa I	RASQSLVSISSYLA (SEQ ID NO: 5)	AASSLES (SEQ ID NO: 6)	QQYNSLPEWT (SEQ ID NO: 7)
Human kappa II	RSSQSLLHSDGDTYLN (SEQ ID NO: 8)	LVSNRAS (SEQ ID NO: 9)	MQALQPRT (SEQ ID NO: 10)
Human kappa III	RASQSVSSSYLA (SEQ ID NO: 11)	GASSRAT (SEQ ID NO: 12)	QQYGSSPPLT (SEQ ID NO: 13)
Human kappa IV	KSSQSVLYSSNNKNYLA (SEQ ID NO: 14)	WASTRES (SEQ ID NO: 15)	QQYYSTPT (SEQ ID NO: 16)
Human lambda I	SGSSSNIIGNNYVS (SEQ ID NO: 17)	DNNKRPS (SEQ ID NO: 18)	ATWDDSLSANSAPV (SEQ ID NO: 19)
Human lambda II	TGTSSDVGGYNAVS (SEQ ID NO: 20)	DVTDRPS (SEQ ID NO: 21)	SSYGGGSNV (SEQ ID NO:22)
Human lambda III	SGDNLGDKYVH (SEQ ID NO: 23)	DDNKRPS (SEQ ID NO:24)	QAWDSSSDHPGVV (SEQ ID NO:25)
Mouse kappa I	KSSQSLLNSGNQKNYLA (SEQ ID NO: 26)	WASTRES (SEQ ID NO:27)	QNDYSYPLT (SEQ ID NO: 28)
Mouse kappa II	RSSQSLVHSNGNTYLE (SEQ ID NO: 29)	KVSNRFS (SEQ ID NO:30)	FQGTHVPPYT (SEQ ID NO: 31)
Mouse kappa III	RASESVDSYGNSFMH (SEQ ID NO: 32)	AASNLES (SEQ ID NO: 33)	QQSNEDPPWT (SEQ ID NO: 34)
Mouse kappa IV	SASSSVSSSYLH (SEQ ID NO: 35)	RTSNLAS (SEQ ID NO: 36)	QQWSSYPGLT (SEQ ID NO: 37)
Mouse kappa V	RASQDDISNYLN (SEQ ID NO: 38)	YASRLHS (SEQ ID NO: 39)	QQGNTLPPRT (SEQ ID NO: 40)
Mouse kappa IV	SASSSVSYMH (SEQ ID NO: 41)	DTSKLAS (SEQ ID NO: 42)	QQWSSNPPMPLT (SEQ ID NO: 43)

analysis of the binding protein with bound ligand or antigen. See, for example, Amit et al., *Science 233*:4765, 747-53, 1986.

Immunoglobulin constant region: a portion of an immunoglobulin polypeptide that follows the carboxy terminus of the variable region. This is usually in the vicinity of amino acid #108 in light chains and amino acid #114 in heavy chains. Constant regions determine the isotype designation of the immunoglobulin and include but are not limited to kappa or lambda light chain constant regions and gamma, mu, alpha, epsilon and delta heavy chain constant regions. Constant regions of heavy chains are divided into domains. The first domain following the variable region is designated CH1. Domains following CH1 include the hinge region, CH2, CH3 and possibly CH4 and membrane spanning domain. Examples of immunoglobulin constant regions can be found in Kabat et al., Sequences of Immunological Interest, National Institutes of Health, Bethesda, Md. 1991.

Tailpiece of a constant region: tailpiece regions are located after the CH3 or CH4 segments of IgA or IgM constant regions, respectively. Examples of tailpieces are listed in Table 3. Other tailpiece regions may be identified based on similarity in sequences to one of the representative heavy chain tailpiece regions in Table 3. In general, a tailpiece should be at least 50% identical to a sequence in Table 3 and always contains a cysteine as the penultimate carboxy terminus amino acid. When co-expressed in appropriate eukaryotic cells containing an endomembrane system, tailpiece regions attached to either alpha or mu constant regions are capable of forming a disulfide bond between the penultimate cysteine of the tailpiece and a cysteine in J chain. Formation of these disulfide bonds can result in the polymerization of alpha or mu constant regions.

Table 3

Representative Heavy Chain Tailpiece Regions

Isotype	Species	Sequence
IgA	Human	PTHVNVSVVMAEVDGTCY (SEQ ID NO: 62)

agent such as dithiothreitol or mercaptoethanol. Comparative analysis of peptides can be accomplished by, for example, denaturing gel electrophoresis using SDS and polyacrylamide as described in Molecular Cloning: A Laboratory Manual, Sambrook et al., eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989) and *Antibodies: A Laboratory Manual*, Harlow and Lane, eds., Cold Spring Harbor, N.Y. (1988).

Alpha chain: a polypeptide that is substantially identical (*i.e.*, at least 90% identical) in sequence to the constant region of an alpha chain of Kabat et al. (Sequences of Immunological Interest, National Institutes of Health, Bethesda, Md. 1991).

Mu chain: a polypeptide that is substantially identical (*i.e.*, at least 90% identical) in sequence to the constant region of a mu chain of Kabat et al. (Sequences of Immunological Interest, National Institutes of Health, Bethesda, Md. 1991).

Secretory component: a polypeptide that binds to polymeric immunoglobulins containing J chains. Secretory components are derived from the polyimmunoglobulin receptor. The sequences of some secretory components and polyimmunoglobulin receptors have been determined. Other polypeptide that share at least 75% sequence identity with a known secretory component, and that retain the ability to bind to polymeric immunoglobulins containing J chains, are also considered to be secretory components.

Heavy chain: a polypeptide that comprises an amino acid sequence that is at least 90% identical to the constant region of a native heavy chain sequence and an amino acid sequence that is at least 75% identical to the variable region of a native heavy chain sequence (*see* Kabat et al., Sequences of Immunological Interest, National Institutes of Health, Bethesda, Md. 1991) and that, when co-expressed with a light chain in the endomembrane system of an appropriate eukaryotic cell is capable of forming a heavy chain-light chain complex, potentially an antibody, joined by disulfide bonding. Heavy chains may be identified as the larger of the two polypeptides present within a divalent antibody, Fab, Fab'2 or Fv. Heavy chains are also found as components of polyvalent antibodies such as IgAs and IgMs.

Dicotyledon (dicot): A flowering plant whose embryos have two seed halves or cotyledons. Examples of dicots are: tobacco; tomato; the legumes including alfalfa; oaks; maples; roses; mints; squashes; daisies; walnuts; cacti; violets; and buttercups.

DNA: Deoxyribonucleic acid.

Epitope: A portion of a molecule that is specifically recognized by an immunoglobulin. It is also referred to as the determinant or antigenic determinant.

Eukaryotic hybrid vector: A DNA molecule by means of which DNA coding for a polypeptide (insert) can be introduced into a eukaryotic cell.

Fab fragment: A polypeptide consisting of a portion of an antibody molecule containing the active portions of an antibody heavy chain and an antibody light chain covalently coupled together and capable of specifically combining with antigen. Fab fragments are typically prepared by proteolytic digestion of substantially intact antibody molecules with papain using methods that are well known in the art. However an Fab fragment may also be prepared by expressing in a suitable host cell the desired portions of an antibody heavy chain and light chain using methods well known in the art.

Fv fragment: A polypeptide consisting of the active portions of an antibody heavy chain variable region and an antibody light chain variable region covalently coupled together and capable of specifically combining with antigen. Fv fragments are typically prepared by expressing in a suitable host cell the desired portions of antibody heavy chain variable region and light chain variable region using methods well known in the art.

Mutagenesis: A process whereby the nucleotide sequence of an original polynucleotide is changed in one or a few locations to produce derivative polynucleotides of substantially the same sequence. Mutagenesis can be accomplished by manipulation of polynucleotides *in vitro* by, for example, using various commonly available enzymes and mutagenic oligonucleotides. Mutagenesis can also be accomplished *in vivo* using the immune system of an animal to introduce desired changes in polynucleotides encoding immunoglobulins for example. B cell maturation,

Polypeptide and peptide: A series of amino acid residues covalently connected one to the other by peptide bonds between the alpha-amino and carboxy groups of adjacent residues or by disulfide bridges between two cysteines.

Protein: A linear series of greater than about 50 amino acid residues connected one to the other as in a polypeptide.

Promoter: A recognition site on a DNA sequence or group of DNA sequences that provide an expression control element for a gene and to which RNA polymerase specifically binds and initiates RNA synthesis (transcription) of that gene. An "inducible promoter" is a promoter where the rate of RNA polymerase binding and initiation is modulated by external stimuli. Such stimuli include light, heat, anaerobic stress, alteration in nutrient conditions, presence or absence of a metabolite, presence of a ligand, microbial attack, wounding and the like. A "viral promoter" is a promoter with a DNA sequence substantially similar to the promoter found at the 5' end of a viral gene. A typical viral promoter is found at the 5' end of the gene coding for the p21 protein of MMTV described by Huang et al., *Cell 27*:245, 1981. A "synthetic promoter" is a promoter that was chemically synthesized rather than biologically derived. Usually synthetic promoters incorporate sequence changes that optimize the efficiency of RNA polymerase initiation. A "constitutive promoter" is a promoter where the rate of RNA polymerase binding and initiation is approximately constant and relatively independent of external stimuli. Examples of constitutive promoters include the cauliflower mosaic virus 35S and 19S promoters described by Poszkowski et al., *EMBO J. 3*:2719, 1989 and Odell et al., *Nature 313*:810, 1985. A "temporally regulated promoter" is a promoter where the rate of RNA polymerase binding and initiation is modulated at a specific time during development. Examples of temporally regulated promoters are given in Chua et al., *Science 244*:174-181, 1989. A "spatially regulated promoter" is a promoter where the rate of RNA polymerase binding and initiation is modulated in a specific structure of the organism such as the leaf, stem or root. Examples of spatially regulated promoters are given in Chua et al., *Science 244*:174-181, 1989. A "spatiotemporally regulated promoter" is a promoter where the rate of RNA polymerase binding and initiation is modulated in a specific structure of

previously exist in the organism. Examples of added traits are resistance to toxic chemicals and expression of foreign proteins not normally produced by the organism.

IMMUNOGLOBULIN BINDING PROTEINS (IGBP) POLYPEPTIDES

As noted above, IgBPs comprise one or more polypeptides in which amino acids are linked by covalent peptide bonds. In general, an IgBP (i) comprises an amino acid sequence that is at least 75% identical to at least one framework region of a native immunoglobulin molecule (*e.g.*, IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda) and (ii) is a functional binding protein. Sequence identity may be determined using any of a variety of well known algorithms, which may be readily optimized by those of ordinary skill in the art. One such algorithm is employed by the Align program described by Dayhoff et al., *Meth. Enzymol. 91*:524-545, 1983. A functional binding protein, as discussed above, specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter (preferably $<10^{-7}$ moles/liter). K_D may be readily determined using well known assays.

An IgBP may comprise a single immunoglobulin chain or fragment thereof, multiple identical immunoglobulin chains or fragments thereof, or multiple non-identical immunoglobulin chains or fragments thereof. IgBPs include, for example, single chain antigen binding proteins, Fabs and Fvs. Other IgBPs are heavy chain binding proteins (C_HBPs), which comprise multiple combining sites composed of amino acid residues derived from the constant region of an immunoglobulin heavy chain and a variable region from any source (*e.g.*, either heavy or light chain, but not both). In a preferred embodiment the variable region is derived from a heavy chain. C_HBPs further comprise an amino acid sequence that is at least 75% identical to a constant region tailpiece of a mu or alpha chain of a native immunoglobulin heavy chain. A C_HBP may further comprise one or more J chains, which can serve to link other component polypeptides. Representative C_HBPs include proteins assembled from four alpha chains and one J chain, from twelve mu chains or from ten mu chains and at least on J chain.

IgBPs may be made up of component polypeptides linked by covalent bonds, preferably disulfide bonds. Preferred components comprise one or more

potential immunoglobulin binding proteins can also be derived from the mutagenesis of unique sequences of immunoglobulins.

Methods for isolating polynucleotides encoding a population of IgBPs are well known in the art. See, for example, Molecular Cloning: A Laboratory Manual, Sambrook et al., eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989) and Antibodies: A Laboratory Manual, Harlow and Lane, eds., Cold Spring Harbor, N.Y. (1988). As noted above, particularly preferred polynucleotides encode immunoglobulin heavy and light chain variable regions, or portions thereof. Such polynucleotides may be isolated from cells obtained from a vertebrate, preferably a mammal, which has been immunized with an antigenic ligand (antigen) against which activity is sought (*i.e.*, a preselected antigen). The immunization can be carried out conventionally and antibody titer in the animal can be monitored to determine the stage of immunization that corresponds to the affinity or avidity desired. Partially immunized animals typically receive only one immunization and cells are collected therefrom shortly after a response is detected. Fully immunized animals display a peak titer that is achieved with one or more repeated injections of the antigen into the host mammal, normally at two to three week intervals. Usually three to five days after the last challenge, the spleen is removed and the genes coding for immunoglobulin heavy and immunoglobulin light chains are isolated from the rearranged B cells present in the spleen using standard procedures. See Current Protocols in Molecular Biology, Ausubel et al., eds., John Wiley and Sons, New York (1987) and Antibodies: A Laboratory Manual, Harlow and Lane, eds., Cold Spring Harbor, N.Y. (1988).

In addition to the spleen, rearranged B cells can be derived from the bone marrow of individuals who have been exposed to specific antigens. For example, patients in hospitals can be exposed to infectious organisms that they would not ordinarily encounter. Frequently, these patients mount an immune response that results in B cell maturation and deposition in bone marrow of B cells expressing antibodies that neutralize the pathogen. In addition to nosocomial exposure, memory B cells are deposited in response to a variety of infections, including HIV, HPV, HSV and CMV.

of choice are introduced. One or a few polypeptides are then expressed for evaluation of binding properties. At the other extreme is random mutagenesis, by means of relatively nonspecific changes of codons at a variety of sites in the gene of choice. Arrays of plants and plant cells, as well as other eukaryotic cells, can be used for functional screens of mutant IgBPs. Preferably, polynucleotides within such arrays encode variants that differ from a native IgBP sequence in one or more amino acid substitutions and/or deletions, such that each variant retains at least 75%, preferably at least 95%, identity to the native IgBP. Particularly preferred are arrays in which each polynucleotide differs from the native IgBP in one or more point mutations. Mutagenesis is also a process that occurs naturally during the development of antibodies. During the course of B cell maturation, antibody-encoding genes are recombined, selected, and mutated to produce plasma cells that encode antibodies with higher affinity for an antigen compared to the original polynucleotide encoding the antibody with specificity for the same antigen.

Regardless of the precise type of IgBP, array polynucleotides are formulated so as to permit entry into, and replication within, a target host cell. For certain transfection techniques, the polynucleotides are cloned into a suitable expression vector. Any vector can be used for such transfection, provided that the vector is capable of the transcription of IgBP genes as well as selectable markers in a target cell. Typical expression vectors useful for expression of genes in plants are well known in the art and include vectors derived from the tumor-inducing (Ti) plasmid of *Agrobacterium tumefasciens* described by Rogers et al., *Meth. in Enzymol. 153*:253-277, 1987. However, several other expression vector systems are known to function in plants. See, for example, Register et al., *Plant Mol Biol. 25*:951-961, 1994; Verch et al., *J. Immunol. Meth. 220*:69-75, 1998.

Alternatively, for certain transfection techniques an expression vector is unnecessary or undesirable. In these cases, either single or multiple DNA fragments containing the polypeptide coding genes linked to plant expression control elements (expression cassettes) are introduced directly into target plant cells. The DNA

A variety of methods have been developed to operatively link DNAs to vectors via complementary cohesive termini. For instance, complementary homopolymer tracks can be added to the DNA segment to be inserted and to the vector DNA. The vector and DNA segment are then joined by hydrogen bonding between the complementary homopolymeric tails to form recombinant DNA molecules. Alternatively, synthetic linkers containing one or more restriction endonuclease sites can be used to join the DNA segment to the expression vector. The synthetic linkers are attached to blunt-ended DNA segments by incubating the blunt-ended DNA segments with a large excess of synthetic linker molecules in the presence of an enzyme that is able to catalyze the ligation of blunt-ended DNA molecules, such as bacteria phage T4 DNA ligase. Thus, the products of the reaction are DNA segments carrying synthetic linker sequences at their ends. These DNA segments are then cleaved with the appropriate restriction endonuclease and ligated into an expression vector that has been cleaved with an enzyme that produces termini compatible with those of the synthetic linker. Synthetic linkers containing a variety of restriction endonuclease sites are commercially available from a number of sources including New England BioLabs, Beverly, MA. Polymerase chain reaction (PCR) can also be used to introduce appropriate restriction sites into polynucleotide sequences as would be needed for ligating the polynucleotide into a restriction site in a plant or eukaryotic cell expression vector.

IMMUNOGLOBULIN BINDING PROTEINS (IGBP) ARRAYS

A library of polynucleotides as described above may be used to generate an array in eukaryotic cells or organisms (*e.g.*, plants or seeds) using standard transfection techniques appropriate for the cell or organism of interest. In general, a library is used to transfect a population of eukaryotic cells such that some or all of the cells contain one or more polynucleotides encoding IgBP polypeptides that are not be detectably expressed by untransfected cells. Such transfection can result in functional transformation, which permits the cells to produce IgBPs and/or components thereof. The cells may then be grown on an appropriate medium to allow for replication and the

In general, recombinant immunoglobulins can be prepared by isolating DNA fragments corresponding to the heavy and light chain variable regions of a monoclonal antibody and joining them to each other by any one of the standard methods known to those of skill in the art and described by Molecular Cloning: A Laboratory Manual, Sambrook et al., eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989). In one preferred embodiment, only DNA fragments corresponding to heavy chains are used to prepare suitable vectors for cell transformation. In one example of cell transformation, the recombinant DNA fragments can be inserted into *Agrobacterium* transfer vectors such that the genes of interest are inserted into the *Agrobacterium* genome. The recombinant *Agrobacterium* is then used to infect plant cells resulting in the production of the polypeptide of interest. In another example of cell transformation, the recombinant DNA fragments can be inserted into baculovirus transfer vectors such that the genes of interest are inserted into the viral genome in lieu of the baculovirus polyhedron gene. The recombinant virus is then used to infect insect cells resulting in the production of the polypeptide of interest.

An advantage of using insect cells that utilize recombinant baculoviruses for the production of IgBPs is that the baculovirus system allows production of mutant antibodies as well as combinatorial expression of immunoglobulin with other polypeptides (such as J chain) much more rapidly than stably transfected mammalian or plant cell lines. In addition, insect cells have been shown to correctly process and glycosylate eukaryotic proteins. Finally, the baculovirus expression of foreign protein has been shown to constitute as much as 50-75% of the total cellular protein late in viral infection, making this system an excellent means of producing milligram quantities of recombinant immunoglobulins.

The use of baculovirus *Autographica californica* nuclear polyhedron virus (AcNPV) and recombinant viral stocks in *Spodoptera frugiperda* (Sf9) cells to prepare large quantities of protein has been described by Smith et al., *Proc. Natl. Acad. Sci. U.S.A. 82*: 8404-8408, 1985, and Summers and Smith, *Bulletin B – Texas Agricultural Experiment Station,* May, 1987. A preferred method of preparing recombinant heavy chains covalently linked to J chain or the light chain constant region

transcripts are then produced (using, for example, the SP6 promoter) containing non-structural genes for *in vivo* replication of the recombinant RNA and the promoter and transgene of interest. The RNA is then used to transfect mammalian or other cells, such as baby hamster kidney (BHK) cells. Protein expression can be assayed 4-72 hours post infection. Alternatively, viral particles can be harvested and used for infection of another cell line.

In certain embodiments, an array comprises plants or cells transformed with polynucleotides encoding immunoglobulin heavy chains and immunoglobulin light chains, such that antibody molecules are produced. Within such arrays, individual immunoglobulin heavy and light chains produced by each plant or plant cell may associate with each other and assume a conformation having an antigen binding site specific for a preselected or predetermined antigen, as evidenced by its ability to be competitively inhibited. When the binding protein is an antigen binding protein, its affinity or avidity is generally greater than 10^5 M^{-1}, preferably greater than 10^6 M^{-1}, and more preferably greater than 10^8 M^{-1}. Immunoglobulins for use in such embodiments may generally be derived from the B cells of an immunized host, each of which express a different IgBP.

Similarly, an IgBP array may comprise polynucleotides that encode portions of immunoglobulin heavy chains and portions of immunoglobulin light chains. The individual immunoglobulin heavy and light chain portions in each plant or plant cell may associate with each other and assume a conformation having an antigen binding site specific for a preselected or predetermined antigen. The antigen binding site on a Fab fragment has a binding affinity or avidity similar to the antigen binding site on an immunoglobulin molecule. Likewise, the antigen binding site on a SCAB protein has a binding affinity or avidity similar to the antigen binding site on an immunoglobulin molecule. Alternatively, an IgBP array may comprise polynucleotides that encode immunoglobulin heavy and light chain fragments that may associate within the plant cell to form Fv fragments with a biologically active conformation that has a binding site specific for a preselected or predetermined antigen. The antigen binding

defined by the border sequences and intervening DNA is usually inserted into the plant genome as described by Spielmann et al., *Mol. Gen. Genet. 205*:34, 1986 and Jorgensen et al., *Mol. Gen. Genet. 207*:471, 1987. Modern *Agrobacterium* transformation vectors are capable of replication in *Escherichia coli* as well as *Agrobacterium*, allowing for convenient manipulations, as described by Klee et al., in *Plant DNA Infectious Agents*, T. Hohn and J. Schell, eds., Springer-Verlag, New York (1985) pp. 179-203. Further recent technological advances in vectors for *Agrobacterium*-mediated gene transfer have improved the arrangement of genes and restriction sites in the vectors to facilitate construction of vectors capable of expressing various polypeptide coding genes. The vectors described by Rogers et al., *Methods in Enzymology 153*:253, 1987, for example, have convenient multi-linker regions flanked by a promoter and a polyadenylation site for direct expression of inserted polypeptide coding genes, and are suitable for present purposes.

In those plant species where *Agrobacterium*-mediated transformation is efficient, this is the method of choice because of the facile and defined nature of the gene transfer (*see* Horsch et al., *Science 227*:1129-1231, 1985; Feldmann and Marks, *Mol. Gen. Genet. 208*:1-9, 1987; Chang et al., *Plant J.* 5:551-558, 1994; Bechtold et al., *Acad. Sci. Paris Science de la Vie 316*:1194-1199, 1993; Hansen and Chilton, *Curr Top Microbiol Immunol 240*:21-57, 1999). Few monocots appear to be natural hosts for *Agrobacterium*. However, transgenic plants may be produced in a variety of monocots using *Agrobacterium* vectors (as described, for example, by Bytebier et al., *Proc. Natl. Acad. Sci. U.S.A. 84*:5345, 1987; Hiei et al., *Plant J. 6(2)*:271-282, 1994; Ishida et al., *Nat. Biotechnol. 14(6)*:745-750, 1996; Hansen and Chilton, *Curr Top Microbiol Immunol 240*:21-57, 1999). Alternatively, vector-free or direct DNA transfer methods have been developed to transform a variety of plant species. As an example, transformation of plant protoplasts can be achieved using methods based on calcium phosphate precipitation, polyethylene glycol treatment, electroporation, microinjection, and combinations of these treatments. See, for example, Potrykus et al., *Mol. Gen. Genet. 199*:183, 1985; Lorz et al., *Mol. Gen. Genet. 199*: 178, 1985; Fromm et al., *Nature 319*:791, 1986; Uchimiya et al., *Mol. Gen. Genet. 204*:204, 1986; Callis et al.,

important crop species such as soybean (McCabe et al., 1988), maize (Fromm et al., *Bio/technology* 8: 833-839,1990; Gordon-Kamm et al, *The Plant Cell 2*: 603-618 1990), rice (Christou et al., *Bio/technology 9*: 957-962, 1991), barley (Wan and Lemaux, *Plant Physiology 104*: 37-48, 1994) and wheat (Vasil et al., *Bio/technology 10*: 662-674, 1992; Weeks et al., *Plant Physiology 102*: 1077-1084, 1993). Transformation of tissue explants eliminates the need for passage through a protoplast stage and thus speeds the production of transgenic plants.

DNA can also be introduced into plants by direct DNA transfer into pollen, as described by Zhou et al., *Methods in Enzymology 101*:433, 1983; Hess, *Intern Rev. Cytol. 107*:367, 1987; Luo et al., *Plant Mol. Biol. Reporter 6*:165, 1988; and Saunders et al., *Molecular Biotechnology 3*:181-190, 1995. Expression of polypeptide coding genes can be obtained by injection of the DNA into reproductive organs of a plant as described by Pena et al., *Nature 325*:274, 1987. DNA can also be injected directly into the cells of immature embryos, followed by the rehydration of desiccated embryos as described by Neuhaus et al., *Theor. Appl. Genet. 75*:30, 1987; and Benbrook et al., in *Proceedings Bio. Expo. 1986*, Butterworth, Stoneham, Mass., pp. 27-54 (1986).

Regardless of the method of transformation, transformed plant cells are generally grown on a suitable medium. Cells may be grown on medium that does not promote differentiation (*e.g.*, as calli), or may be grown such that plants are regenerated. For example, some lower plant species, such as *Chlamydomonas*, do not require regeneration and can express immunoglobulin binding proteins immediately after selection for the expression of a selectable marker, as a result of growth in an appropriate medium. Suitable media for a wide variety of plant cell types are well known to those of ordinary skill in the art. In general, transformed cells are selected based on expression of the selection marker, such that only those cells that contain at least one transfected polynucleotide encoding an immunoglobulin binding protein are permitted to grow.

Non-combinatorial arrays expressing heavy chain binding proteins may be generated using, for example, the MaxBac Baculovirus Expression System

expression may occur during seed development). Nonetheless, such a set of seeds is considered an array.

Within arrays, the IgBPs or components thereof may be secreted from cells or may accumulate in one or more intracellular compartments. Secreted proteins may or may not be contained by a cell wall. For example, the CW15 mutant of *Chlamydomonas* does not express a cell wall and IgBPs containing leader sequences are secreted directly into the growth medium. In this instance, the desired IgBP can be identified by assaying a portion of the growth medium for the desired characteristic. In higher plants, secreted IgBPs are generally contained by the cell wall and may accumulate in the apoplastic water. A desired IgBP can be assayed after disruption of the cell wall using any of a variety of standard mechanical techniques (*e.g.*, mortar and pestle homogenization).

Within certain embodiments, an array of IgBPs in plant cells is derived from a progeny population of plant cells (*i.e.*, plant cells resulting from the sexual cross of transformants expressing IgBP components). In such an array, some or all of the progeny population contain polynucleotides encoding IgBP components, such that functional IgBPs are expressed in the array. In such arrays, useful IgBP component polynucleotides include those that encode a second polypeptide that can autogenously associate with a first polypeptide in such a way as to form a biologically functional IgBP. Examples of such IgBP components are heavy and light chains of antibodies.

The sexual cross of different members of a plant species has been well described by Mendel in 1865 (an English translation of Mendel's original paper together with comments and a bibliography of Mendel by others can be found in Experiments In Plant Hybridization, Edinburgh, Scotland, Oliver Boyd, eds., 1965). When the plants are flowering plants, the sexual cross involves contacting viable pollen from one population of plants with the stigma of another population of plants of a sexually compatible species. When the plant cells are photosynthetic unicellular organisms (*e.g., Chlamydomonas*), the sexual cross involves fusion of cells of + and – mating types, followed by meiosis and tetrad formation. Progeny from sexual crosses can comprise seeds or tetrads.

composition comprises no more than 10,000 grams of plant material for each milligram of IgBP, but at least 100 nanograms of plant material for each gram of IgBP and more preferably at least one milligram of plant material for each 500 milligrams of IgBP.

A composition may further comprise substances such as chlorophyll, synergistic compounds, medicines, compounds derived from medicinal plants and/or various pharmaceuticals.

ASSAYS FOR SCREENING ARRAYS OF IMMUNOGLOBULIN BINDING PROTEINS

Transgenic arrays as provided herein are useful for the discovery of IgBPs having desired characteristics. From the transgenic arrays provided herein, individual eukaryotic organisms or clones of eukaryotic cells can be identified rapidly, enabling the easy access to an economical, high yield process for the large scale production of a desired IgBP.

As noted above, plants are a particularly preferred host organism for the transgenic arrays described herein. When a sexual cross is performed between two populations of plant cells, each expressing IgBPs or components of IgBPs, some or all of the progeny can express functional IgBPs that comprise multiple IgBP components. Such progeny may be of particular use within functional screens for IgBPs having multiple polypeptide subunits.

More specifically, once regenerated plants are adapted to growth in soil, small sections of leaf can be removed to assay for the presence of IgBPs. Alternatively, the plants can be allowed to set seed by either self-pollination or cross pollination and the assays for expression of IgBPs can be performed on individual seeds or portions of seeds. When the transformants are individual cells (*e.g., Chlamydomonas*) the cells may be grown in solid or liquid medium to obtain a sufficient quantity for performing assays. Assays can be immunological (*e.g.*, ELISA) or functional. Functional assays may involve, for example, binding of the IgBP to cognate factors immobilized on a solid support. When the IgBP is an antibody, the cognate factor is an antigen. Assays for determining the number of copies of IgBP as well as the number of different IgBPs being expressed in the primary transformants include Southern blotting and DNA

EXAMPLES

Example 1

Isolation and Expression in Maize of Immunoglobulin Heavy Chain-Coding Genes And Immunoglobulin Light Chain-Coding Genes Derived From B Cells in Bone Marrow

This Example illustrates the preparation of representative immunoglobulin binding protein components for use in preparing an array of immunoglobulin binding proteins in plant cells.

Bone marrow (~10^7 cells) was obtained from donors at local hospitals. Donors were chosen on the basis of (a) having had *Clostridium difficile*-associated diarrhea within the previous two months and (b) having toxin neutralizing antibodies in their serum. Bone marrow was obtained from the subset of this population whose diarrhea resolved without the intervention of any further antibiotic therapy.

Bone marrow cells were washed multiple times in cold PBS then incubated with combined FITC-labeled *Clostridium difficile* toxin A and toxin B. This procedure was used to identify B cells expressing anti-toxin A or anti-toxin B on their cell surfaces. The *Clostridium difficile* toxins were prepared from cultures of *Clostridium difficile* as described by Lyerly et al., *Infect Immun 35*:1147-50, 1982.

Fluorescent labeling of the toxins was by reaction with fluorescein isothiocyanate (FITC) using procedures provided by the manufacturer (Pierce). Incubation of bone marrow cells with the combined FITC-labeled *Clostridium difficile* toxin A and toxin B was by the procedures described in Hoven et al., *J. Immunol. Methods 117*:2275-84, 1989. The tagged cells were shown to be B cells by double staining with phycoerythrin-conjugated anti-human Ig.

Total RNA from the selected bone marrow cells was prepared using the QIAamp RNA Blood Mini-Kit system (Qiagen) according to the instructions provided.

PCR amplification of the isolated RNA employed oligonucleotides complementary to conserved 3' and 5' regions of immunoglobulin (IgG1) transcripts as previously described in Antibody Engineering, Carl A, Borrebaeck, Ed., W.H. Freeman and Company, New York, N.Y., 1995. The oligonucleotide primers introduced unique

The recombinant population of vectors was then used to transform *Agrobacterium tumefasciens* as described by Ishida et al., *Nature Biotechnology 14*:745-75, 1996. The transformed *Agrobacterium* was then used to infect immature corn embryos followed by selection and regeneration of maize plants as described in Ishida et al. The resulting array of plant cells producing IgBP polypeptides (*i.e.*, IgBPs and/or IgBP components) comprised approximately 1000 plants.

To locate the position in the array of various IgBP polypeptides, samples of leaf tissue were taken from each plant. A variety of assays were performed to characterize the presence of IgBPs, the functionality of IgBPs, and the variety of IgBP components in each transformant. IgBP components (gamma chains or kappa chains) were detected by ELISA in 96 well plates. In one set of plates the capture of IgBP components was by reaction with immobilized goat anti-human kappa chain antibody and the detection was by horse radish peroxidase conjugated goat anti-human kappa chain antibody. In another set of plates the capture of IgBP components was by reaction with immobilized goat anti-human gamma chain antibody and the detection was by horse radish peroxidase conjugated goat anti-human gamma chain antibody. In a third set of plates the capture of IgBP components was by reaction with immobilized goat anti-human gamma chain antibody and the detection was by horseradish peroxidase-conjugated goat anti-human kappa chain antibody. Detection of captured IgBP components was by the procedures previously described (Antibodies: A Laboratory Manual, Harlow and Lane, eds., Cold Spring Harbor, N.Y., 1988). The results demonstrated that the majority of plants co-expressed gamma and kappa chains..

The leaf material was further evaluated for the ability of the expressed IgBPs to bind the toxin A or toxin B proteins. In these assays, ELISA plates were coated with either toxin A or B to capture the IgBPs. The ELISA procedures generally followed those previously described (Hiatt et al., *Nature 342*:76-78, 1989). Leaf extracts were incubated in the ELISA wells to allow reaction of potential immunoglobulin binding proteins with the immobilized toxins. Screening the entire array of approximately 1000 plants as described above identified 10 plants expressing functional anti-toxin A antibodies and 5 plants expressing functional anti-toxin B

the progeny. This observation is consistent with the known ability of *Agrobacterium* to introduce more than one vector into a nucleus during transformation.

Example 2

Isolation and Expression in Maize of Immunoglobulin Heavy Chain-Coding Genes And J Chain-Coding Gene

This Example illustrates the preparation of representative immunoglobulin heavy chain binding protein components for use in generating an array of immunoglobulin heavy chain binding proteins in plant cells.

RNA was prepared as described in Example 1. PCR amplification of the isolated RNA employed oligonucleotides complementary to conserved 3' and 5' regions of the variable region of the immunoglobulin heavy chain (IgG1) transcripts as previously described in Antibody Engineering, Carl A, Borrebaeck, Ed., W.H. Freeman and Company, New York, N.Y., 1995. The oligonucleotide primers introduced unique restriction sites at each end of the amplified polynucleotide to allow for ligation into a plant expression vector. The amplified portion of the immunoglobulin polynucleotides extended from the 5' end encoding the respective FR1 regions to the 3' end encoding the first six amino acids of the CH1 region. In addition, appropriate 5' and 3' restriction sites were introduced in the process of amplification. The amplified heavy chain polynucleotide chains therefore corresponded to the first six codons at the 5' end of the RNAs and the first six codons of the constant regions with the restriction sites introduced outside of the respective terminal codons. This immunoglobulin heavy chain binding protein array consisting of double stranded amplified cDNAs from the sorted cells was prepared for insertion into a plant expression vector, by converting the ends of the double stranded cDNA to the appropriate sticky ends for ligation into the vector (Molecular Cloning: A Laboratory Manual, Sambrook et al., eds., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 1989) by digestion with a restriction endonuclease according to the manufacturer's protocol (New England BioLabs). Synthetic polynucleotides encoding J chain were additionally prepared for ligation into the plant vector by PCR amplification as above. Finally, the constant region of an IgG-

as described above where leaf extracts were incubated in the ELISA wells to allow reaction of potential heavy chain binding proteins with the immobilized toxins. Screening the entire array of approximately 1000 plants as described above identified 50 plants expressing functional binders of toxin A and 40 plants expressing functional binders of toxin B. From these results, it appears that the non-combinatorial approach to isolating binding proteins containing no light chain variable regions yields a higher proportion of transformation events expressing binding proteins compared to the combinatorial approach involving the random pairing of heavy and light chains. The ability to detect these binding proteins appears to be due to their assembly into a polymeric molecular structure with enhanced avidity for the toxins.

Example 3

Preparation of an Array of Functional Immunoglobulin Binding Proteins from Binding Protein Components by a Sexual Cross of Higher Plants

This Example illustrates the generation of an array of plants that express functional immunoglobulin binding proteins from plant populations that express different immunoglobulin binding protein components.

cDNAs encoding heavy and light chains derived from the B cell populations described above were cloned separately into vectors allowing for expression of heavy chains and light chains in different plant populations. The procedures described above were performed to generate plant transformants expressing arrays of heavy chains or light chains (immunoglobulin binding protein components). In order to form an array of immunoglobulin binding proteins, pollen from the light chain transformants was used to pollinate the heavy chain transformants to produce a seed set. The seeds were then grown into individual plantlets and subjected to the analysis for identification of immunoglobulin binding proteins described above. Results similar to Example 1 were obtained indicating that the random assortment of heavy and light chains by cloning them into different vectors followed by assembly upon crossing plants is comparable to cloning heavy and light chains into the same vector followed by assembly in primary transformants.

introduced unique Sac and Xho restriction sites at the ends of each DNA. Likewise, a single kappa constant region and gamma constant region (sequence derived from Kabat et al.) were synthesized with *Chlamydomonas* codon bias. The synthetic kappa and gamma constant region DNAs were first restricted and ligated into separate Xho-EcoR1 restricted pBluescript vectors to create separate vectors containing synthetic immunoglobulin constant region polynucleotides. These vectors were cloned and purified and were then used as the recipients for Sac-Xho restricted synthetic variable region DNAs to create an array of separate gamma and kappa polynucleotides encoding the diverse variable regions and constant regions of full length gamma and kappa-encoding polynucleotides. These vectors were used to transform bacteria from which plasmids were purified and automated sequence analysis was used to identify the original 100 variable regions comprising the arrays. These plasmids were mixed and the Sac-EcoR1 restriction fragments were subcloned into the pARG7 vector. The pARG7 vector contained a signal sequence derived from the Kabat database (MDWTGRFLFVVAAATGVQS; SEQ ID NO: 73) upstream from Sac-EcoR1 sites. The ligation product created full length heavy and light chain polynucleotides capable of co-expression with a signal sequence to direct polypeptide synthesis to the endomembrane system of the cell. Colonies on agar plates were tested for expression of antibodies using a nitrocellulose lift assay. Nitrocellulose circles were placed on the agar plates containing colonies on order to adsorb sufficient antibody onto the paper to allow for functional detection.

Example 5

Expression of Heavy Chain Binding Protein Genes in Insect Cells

This example illustrates the generation of an array of insect cells that express functional heavy chain binding proteins.

The heavy chain and J chain PCR products derived from B cells as described above were additionally manipulated to allow for expression in insect cells. In general, the double stranded cDNA from the selected B cells described above was prepared for insertion into an insect expression vector, by converting the ends of the

recommended by the manufacturer to generate an array of viral plaques on agar plates. Each plaque potentially contained the virions encoding unique heavy chain binding proteins as well as the unique heavy chain binding proteins themselves. Virus particles from approximately 1000 of the plaques were used to infect High Five™ insect cells (Invitrogen) in an array of 96 well plates (~10^6 cells per well). 72 hours post-infection cells and supernatant were harvested, brought to 1 mL total volume, and 10-100 µL were analyzed for composition and functionality. Identification of desired heavy chain and binding functionality can utilize any of a variety of techniques known to those of skill in the art such as ELISA, immunoblotting, Western blotting, immunoprecipitation and such, all of which are described in Harlow and Lane, "Using Antibodies: A Laboratory Manual", Cold Spring Harbor Press, Cold Spring Harbor, New York, 1999.

Analysis of the antigen binding capability of each supernatant by ELISA identified a population of heavy chain binding proteins recognizing either toxin A or toxin B. Approximately 5% of the wells containing transformed insect cells expressed detectible binding proteins recognizing toxin A and 3% expressed binding proteins recognizing toxin B. These results are similar to those of Example 2.

Example 6

Expression of Heavy Chain Binding Protein Genes in Mammalian Cells

This example illustrates the generation of an array of mammalian cells that express functional heavy chain binding proteins.

The heavy chain and J chain PCR products derived from B cells as described above were additionally manipulated to allow for expression in baby hamster kidney (BHK) cells. In general, the double stranded cDNAs from the selected B cells described in Example 1 was engineered to be expressed as human mu isotypes as described in Example 6 and was further prepared for insertion into the Sindbis expression vector, using methods describe above. The vector (pSinRep5, Invitrogen) was therefore engineered to allowed for the expression of heavy and J chains using P_{SG} promoters and appropriate polyadenylation sequences. The vector therefore was potentially capable of co-expression of both a heavy chain and a J chain.

Claims

1. A method for preparing an immunoglobulin binding protein array in plant cells, comprising the steps of:

(a) transforming a population of plant cells with a library of at least two different polynucleotides encoding different immunoglobulin binding protein (IgBP) polypeptides that:

(i) specifically bind to a ligand with a $K_D<10^{-6}$ moles/liter; or

(ii) form one or more disulfide bonds with one or more polypeptides in the transfected cell, to generate a binding protein that specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter;

wherein the IgBP polypeptides (i) comprise four framework regions alternating with three complementarity determining regions and (ii) comprise at least one peptide sequence having at least 75% sequence identity to a framework region of a native IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda immunoglobulin molecule; and wherein the IgBP polypeptides are not detectably expressed by the plant cells prior to transformation; and

(b) selecting transformed plant cells, and therefrom preparing an IgBP array in plant cells.

2. A method according to claim 1, wherein each IgBP polypeptide is a functional IgBP.

3. A method according to claim 1, wherein each IgBP polypeptide is an IgBP component that, upon disulfide linkage to one or more IgBP components encoded by other polynucleotides in the library, forms a functional IgBP.

4. A method according to claim 1, further comprising the step of:

consisting of heavy chains and fragments thereof, light chains and fragments thereof, J chains and secretory components.

13. A method according to claim 1, wherein the polypeptides retain at least 95% amino acid identity to a framework region of a native IgM, IgG, IgA, IgD, IgE, IgY, kappa or lambda immunoglobulin molecule.

14. A method according to claim 1, wherein the framework regions are human.

15. A method according to claim 1, wherein the framework regions are murine.

16. A method according to claim 1, wherein the step of transforming is performed via *Agrobacterium*-mediated transformation, chemically-induced DNA uptake, electroporation, solid particle intrusion, biolistics, microinjection, macroinjection, lipofection or viral infection.

17. A method according to claim 1, wherein the IgBP polypeptides are secreted from the plant cells.

18. A method according to claim 1, wherein the plant cells are dicotyledonous plant cells.

19. A method according to claim 18, wherein the plant cells are tobacco or *Arabidopsis* plant cells.

20. A method according to claim 1, wherein the plant cells are monocotyledonous plant cells.

21. A method according to claim 20, wherein the plant cells are corn, Lemna or rice plant cells.

26. A method according to claim 25, wherein the polynucleotides encode immunoglobulin alpha or mu chains.

27. A method according to claim 25, wherein the cells are further transformed with one or more polynucleotides encoding polypeptides having sequences that are at least 75% identical to a sequence of an immunoglobulin J chain.

28. A method according to claim 25, wherein each C_HBP is assembled from four alpha chains and one J chain.

29. A method according to claim 25, wherein each C_HBP is assembled from twelve mu chains.

30. A method according to claim 25, wherein each C_HBP is assembled from ten mu chain and at least one J chain.

31. A method according to claim 25, wherein the C_HBPs or components thereof further comprise one or more portions of immunoglobulin molecules selected from the group consisting of J chains, secretory components and light chain constant regions.

32. A method according to claim 25, wherein the cells are plant cells.

33. A method according to claim 25, wherein the cells are insect cells.

34. A method according to claim 25, wherein the cells are mammalian cells.

39. A method according to claim 25 or claim 35, wherein the library comprises at least 100 different polynucleotides.

40. A method according to claim 25 or claim 35, wherein the library comprises at least 1000 different polynucleotides.

41. A method according to claim 25 or claim 35, wherein the library comprises at least 10,000 different polynucleotides.

42. A method according to claim 25 or claim 35, wherein the polynucleotides encode polypeptides that retain at least 95% amino acid identity to a contant region tailpiece of a mu or alpha chain of a native immunoglobulin heavy chain.

43. A method according to claim 25 or claim 35, wherein the step of transforming is performed via *Agrobacterium*-mediated transformation, chemically-induced DNA uptake, electroporation, solid particle intrusion, biolistics, microinjection, macroinjection, lipofection or viral infection.

44. A method according to claim 25 or claim 35, wherein the binding proteins accumulate in an intracellular compartment of the cells.

45. A method according to claim 25 or claim 35, wherein the binding proteins are secreted from the cells.

46. A method according to claim 32 or claim 35, wherein the plant cells are dicotyledonous plant cells.

47. A method according to claim 46, wherein the plant cells are tobacco or Arabidopsis plant cells.

wherein the cells assemble C_HBPs comprising at least four combining sites.

54. A binding protein array according to claim 53, wherein the polynucleotides encode polypeptide components of immunoglobulin molecules independently selected from the group consisting of heavy chains and fragments thereof, light chains and fragments thereof, J chains and secretory components.

55. A binding protein array according to claim 53, wherein the cells are plant cells.

56. A binding protein array according to claim 53, wherein the cells are insect cells.

57. A binding protein array according to claim 53, wherein the cells are mammalian cells.

58. A binding protein array according to claim 55, wherein the plant cells are selected from the group consisting of corn, rice, Lemna, tobacco and *Chlamydomonas*.

59. A binding protein array according to claim 53, wherein at least 10 different binding proteins are assembled by the cells in the array.

60. A binding protein array according to claim 53, wherein at least 100 different binding proteins are assembled by the cells in the array.

61. A binding protein array according to claim 53, wherein at least 100 different binding proteins are assembled by the cells in the array.

(b) transforming a population of eukaryotic cells with the library of heavy chain variants; and

(c) growing the transformed cells on a medium that permits assembly of C_HBPs, wherein each C_HBP comprises at least four combining sites;

and therefrom preparing a C_HBP array in eukaryotic cells.

66. A C_HBP that:

(a) comprises an amino acid sequence that is at least 75% identical to a constant region tailpiece of a mu or alpha chain of a native immunoglobulin heavy chain;

(b) comprises at least four combining sites, wherein all of the combining sites satisfy the same one of the following requirements:

(i) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin light chain variable region; or

(ii) at least 75% identity to a 25 consecutive amino acid portion of an immunoglobulin heavy chain variable region; and

(c) either (i) specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter; or (ii) forms one or more covalent bonds with one or more polypeptides in a cell, to generate a C_HBP that specifically binds to a ligand with a $K_D<10^{-6}$ moles/liter.

<212> PRT
<213> Homo sapien - Mus musculus

<220>
<221> VARIANT
<222> (1)...(4)
<223> Xaa = Any Amino Acid

<400> 2
Cys Xaa Xaa Xaa
1

<210> 3
<211> 5
<212> PRT
<213> Homo sapien - Mus musculus

<400> 3
Leu Glu Trp Ile Gly
1 5

<210> 4
<211> 4
<212> PRT
<213> Homo sapien - Mus musculus

<220>
<221> VARIANT
<222> (1)...(4)
<223> Xaa = Any Amino Acid

<400> 4
Trp Gly Xaa Gly
1

<210> 9
<211> 7
<212> PRT
<213> Homo sapien

<400> 9
Leu Val Ser Asn Arg Ala Ser
1 5

<210> 10
<211> 8
<212> PRT
<213> Homo sapien

<400> 10
Met Gln Ala Leu Gln Pro Arg Thr
1 5

<210> 11
<211> 12
<212> PRT
<213> Homo sapien

<400> 11
Arg Ala Ser Gln Ser Val Ser Ser Ser Tyr Leu Ala
1 5 10

<210> 12
<211> 7
<212> PRT
<213> Homo sapien

<400> 12
Gly Ala Ser Ser Arg Ala Thr

<400> 16
Gln Gln Tyr Tyr Ser Thr Pro Thr
1 5

<210> 17
<211> 14
<212> PRT
<213> Homo sapien

<400> 17
Ser Gly Ser Ser Ser Asn Ile Ile Gly Asn Asn Tyr Val Ser
1 5 10

<210> 18
<211> 7
<212> PRT
<213> Homo sapien

<400> 18
Asp Asn Asn Lys Arg Pro Ser
1 5

<210> 19
<211> 14
<212> PRT
<213> Homo sapien

<400> 19
Ala Thr Trp Asp Asp Ser Leu Ser Ala Asn Ser Ala Pro Val
1 5 10

<210> 20
<211> 14
<212> PRT

<212> PRT
<213> Homo sapien

<400> 24
Asp Asp Asn Lys Arg Pro Ser
1 5

<210> 25
<211> 13
<212> PRT
<213> Homo sapien

<400> 25
Gln Ala Trp Asp Ser Ser Ser Asp His Pro Gly Val Val
1 5 10

<210> 26
<211> 17
<212> PRT
<213> Mus musculus

<400> 26
Lys Ser Ser Gln Ser Leu Leu Asn Ser Gly Asn Gln Lys Asn Tyr Leu
1 5 10 15
Ala

<210> 27
<211> 7
<212> PRT
<213> Mus musculus

<400> 27
Trp Ala Ser Thr Arg Glu Ser
1 5

1 5 10

<210> 32
<211> 15
<212> PRT
<213> Mus musculus

<400> 32
Arg Ala Ser Glu Ser Val Asp Ser Tyr Gly Asn Ser Phe Met His
1 5 10 15

<210> 33
<211> 7
<212> PRT
<213> Mus musculus

<400> 33
Ala Ala Ser Asn Leu Glu Ser
1 5

<210> 34
<211> 10
<212> PRT
<213> Mus musculus

<400> 34
Gln Gln Ser Asn Glu Asp Pro Pro Trp Thr
1 5 10

<210> 35
<211> 12
<212> PRT
<213> Mus musculus

<400> 35

<400> 39
Tyr Ala Ser Arg Leu His Ser
1 5

<210> 40
<211> 10
<212> PRT
<213> Mus musculus

<400> 40
Gln Gln Gly Asn Thr Leu Pro Pro Arg Thr
1 5 10

<210> 41
<211> 10
<212> PRT
<213> Mus musculus

<400> 41
Ser Ala Ser Ser Ser Val Ser Tyr Met His
1 5 10

<210> 42
<211> 7
<212> PRT
<213> Mus musculus

<400> 42
Asp Thr Ser Lys Leu Ala Ser
1 5

<210> 43
<211> 12
<212> PRT
<213> Mus musculus

<210> 47
<211> 7
<212> PRT
<213> Homo sapien

<400> 47
Ser Tyr Gly Trp Ser Trp Asn
1 5

<210> 48
<211> 19
<212> PRT
<213> Homo sapien

<400> 48
Arg Ile Tyr Tyr Arg Ala Tyr Ser Gly Ser Thr Thr Tyr Asn Pro Ser
1 5 10 15
Leu Lys Ser

<210> 49
<211> 19
<212> PRT
<213> Homo sapien

<400> 49
Glu Leu Pro Gly Gly Tyr Thr Gly Asp Asp Tyr Tyr Tyr Gly Ser Gly
1 5 10 15
Phe Asp Val

<210> 50
<211> 5
<212> PRT

1 5

<210> 54
<211> 17
<212> PRT
<213> Mus musculus

<400> 54
Tyr Ile Ser Gly Tyr Ser Gly Ser Thr Tyr Tyr Asn Pro Ser Leu Lys
1 5 10 15
Ser

<210> 55
<211> 18
<212> PRT
<213> Mus musculus

<400> 55
Gly Gly Tyr Gly Tyr Gly Tyr Tyr Tyr Tyr Asp Tyr Tyr Tyr Tyr Phe
1 5 10 15
Asp Tyr

<210> 56
<211> 6
<212> PRT
<213> Mus musculus

<400> 56
Asp Tyr Tyr Met Asn Asn
1 5

<210> 57
<211> 17

Val Lys Gly

<210> 61
<211> 19
<212> PRT
<213> Mus musculus

<400> 61
Asp Tyr Tyr Tyr Gly Ser Ser Tyr Tyr Glu Gly Pro Val Tyr Trp Tyr
1 5 10 15
Phe Asp Val

<210> 62
<211> 18
<212> PRT
<213> Homo sapien

<400> 62
Pro Thr His Val Asn Val Ser Val Val Met Ala Glu Val Asp Gly Thr
1 5 10 15
Cys Tyr

<210> 63
<211> 15
<212> PRT
<213> Mus musculus

<400> 63
Pro Thr Val Asn Val Ile Met Ser Glu Gly Asp Gly Ile Cys Tyr
1 5 10 15

<210> 64

<213> Homo sapien

<400> 67
Pro Thr Leu Tyr Asn Val Ser Leu Val Met Ser Asp Thr Ala Gly Thr
1 5 10 15
Cys Tyr

<210> 68
<211> 18
<212> PRT
<213> Homo sapien

<400> 68
Pro Thr Leu Tyr Asn Val Ser Leu Ile Met Ser Asp Thr Gly Gly Thr
1 5 10 15
Cys Tyr

<210> 69
<211> 18
<212> PRT
<213> Cricetinae gen. sp.

<400> 69
Pro Thr Leu Tyr Asn Val Ser Leu Ile Met Ser Asp Ala Gly Gly Thr
1 5 10 15
Cys Tyr

<210> 70
<211> 18
<212> PRT
<213> Cricetinae gen. sp.

<400> 73

Met Asp Trp Thr Gly Arg Phe Leu Phe Val Val Ala Ala Ala Thr Gly
1 5 10 15
Val Gln Ser

<210> 74
<211> 23
<212> PRT
<213> Homo sapien

<400> 74

Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
1 5 10 15
Asp Arg Val Thr Ile Thr Cys
20

<210> 75
<211> 15
<212> PRT
<213> Homo sapien

<400> 75

Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr
1 5 10 15

<210> 76
<211> 32
<212> PRT
<213> Homo sapien

<400> 76

Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
1 5 10 15

<213> Homo sapien

<400> 80
Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
1 5 10 15
Leu Lys Ile Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys
20 25 30

<210> 81
<211> 10
<212> PRT
<213> Homo sapien

<400> 81
Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
1 5 10

<210> 82
<211> 23
<212> PRT
<213> Homo sapien

<400> 82
Glu Ile Val Leu Thr Gln Ser Pro Gly Thr Leu Ser Leu Ser Pro Gly
1 5 10 15
Glu Arg Ala Thr Leu Ser Cys
20

<210> 83
<211> 15
<212> PRT
<213> Homo sapien

<400> 83
Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu Ile Tyr

<212> PRT
<213> Homo sapien

<400> 87
Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr
1 5 10 15

<210> 88
<211> 32
<212> PRT
<213> Homo sapien

<400> 88
Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
1 5 10 15
Leu Thr Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys
20 25 30

<210> 89
<211> 10
<212> PRT
<213> Homo sapien

<400> 89
Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
1 5 10

<210> 90
<211> 22
<212> PRT
<213> Homo sapien

<400> 90
Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Gly Ala Pro Gly Gln
1 5 10 15

<213> Homo sapien

<400> 94
Gln Ser Ala Leu Thr Gln Pro Ala Ser Val Ser Gly Ser Pro Gly Gln
1 5 10 15
Ser Ile Thr Ile Ser Cys
20

<210> 95
<211> 15
<212> PRT
<213> Homo sapien

<400> 95
Trp Tyr Gln Gln His Pro Gly Lys Ala Pro Lys Leu Leu Ile Tyr
1 5 10 15

<210> 96
<211> 32
<212> PRT
<213> Homo sapien

<400> 96
Gly Val Pro Asp Arg Phe Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser
1 5 10 15
Leu Thr Ile Ser Gly Leu Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys
20 25 30

<210> 97
<211> 11
<212> PRT
<213> Homo sapien

<400> 97
Phe Gly Gly Gly Thr Lys Val Thr Val Leu Gly

<212> PRT
<213> Homo sapien

<400> 101
Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
1 5 10

<210> 102
<211> 21
<212> PRT
<213> Homo sapien

<400> 102
Ser Glu Leu Thr Gln Pro Pro Ser Val Ser Val Ala Leu Gly Gln Thr
1 5 10 15
Val Arg Ile Thr Cys
20

<210> 103
<211> 15
<212> PRT
<213> Homo sapien

<400> 103
Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Leu Leu Val Ile Tyr
1 5 10 15

<210> 104
<211> 32
<212> PRT
<213> Homo sapien

<400> 104
Gly Ile Pro Asp Arg Phe Ser Gly Ser Ser Ser Gly His Thr Ala Ser
1 5 10 15

<213> Homo sapien

<400> 108
Gly Val Pro Asp Arg Phe Ser Gly Ser Lys Ser Asp Gln Thr Ala Ser
1 5 10 15
Leu Thr Val Ser Gly Leu Arg Ala Glu Asp Glu Ala Asp Tyr Tyr Cys
20 25 30

<210> 109
<211> 11
<212> PRT
<213> Homo sapien

<400> 109
Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Arg
1 5 10

<210> 110
<211> 22
<212> PRT
<213> Homo sapien

<400> 110
Asn Phe Met Leu Thr Gln Pro His Ser Val Ser Glu Ser Pro Gly Lys
1 5 10 15
Thr Val Thr Ile Ser Cys
20

<210> 111
<211> 15
<212> PRT
<213> Homo sapien

<400> 111
Trp Tyr Gln Gln Arg Pro Gly Ser Ala Pro Thr Thr Val Ile Tyr

<212> PRT
<213> Mus musculus

<400> 115
Trp Tyr Gln Gln Lys Pro Gly Gln Ser Pro Lys Leu Leu Ile Tyr
1 5 10 15

<210> 116
<211> 32
<212> PRT
<213> Mus musculus

<400> 116
Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Ser Gly Thr Asp Phe Thr
1 5 10 15
Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Tyr Cys
20 25 30

<210> 117
<211> 10
<212> PRT
<213> Mus musculus

<400> 117
Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
1 5 10

<210> 118
<211> 23
<212> PRT
<213> Mus musculus

<400> 118
Asp Val Val Met Thr Gln Thr Pro Leu Ser Leu Pro Val Ser Leu Gly
1 5 10 15

<213> Mus musculus

<400> 122
Asp Ile Val Leu Thr Gln Ser Pro Ala Ser Leu Ala Val Ser Leu Gly
1 5 10 15
Gln Arg Ala Thr Ile Ser Cys
20

<210> 123
<211> 15
<212> PRT
<213> Mus musculus

<400> 123
Trp Tyr Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Tyr
1 5 10 15

<210> 124
<211> 32
<212> PRT
<213> Mus musculus

<400> 124
Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr
1 5 10 15
Leu Asn Ile His Pro Val Glu Glu Asp Asp Ala Ala Thr Tyr Tyr Cys
20 25 30

<210> 125
<211> 10
<212> PRT
<213> Mus musculus

<400> 125
Phe Gly Gly Gly Thr Lys Leu Glu Ile Lys

<212> PRT
<213> Mus musculus

<400> 129
Phe Gly Ala Gly Thr Lys Leu Glu Ile Lys
1 5 10

<210> 130
<211> 23
<212> PRT
<213> Mus musculus

<400> 130
Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Leu Gly
1 5 10 15
Asp Arg Val Thr Ile Thr Cys
20

<210> 131
<211> 15
<212> PRT
<213> Mus musculus

<400> 131
Trp Tyr Gln Gln Lys Pro Gly Gly Ser Pro Lys Leu Leu Ile Tyr
1 5 10 15

<210> 132
<211> 32
<212> PRT
<213> Mus musculus

<400> 132
Gly Val Pro Ser Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Ser
1 5 10 15

<213> Mus musculus

<400> 136
Gly Val Pro Ala Arg Phe Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser
1 5 10 15
Leu Thr Ile Ser Ser Met Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys
20 25 30

<210> 137
<211> 10
<212> PRT
<213> Mus musculus

<400> 137
Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys
1 5 10

<210> 138
<211> 22
<212> PRT
<213> Mus musculus

<400> 138
Gln Ala Val Val Thr Gln Glu Ser Ala Leu Thr Thr Ser Pro Gly Glu
1 5 10 15
Thr Val Thr Leu Thr Cys
20

<210> 139
<211> 15
<212> PRT
<213> Mus musculus

<400> 139
Trp Val Gln Glu Lys Pro Asp His Leu Phe Thr Gly Leu Ile Gly

<212> PRT
<213> Homo sapien

<400> 143
Trp Val Arg Gly Ala Pro Gly Gln Gly Leu Glu Trp Met Gly
1 5 10

<210> 144
<211> 32
<212> PRT
<213> Homo sapien

<400> 144
Arg Val Thr Ile Thr Ala Asp Thr Ser Thr Ser Thr Ala Tyr Met Glu
1 5 10 15
Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg
20 25 30

<210> 145
<211> 11
<212> PRT
<213> Homo sapien

<400> 145
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
1 5 10

<210> 146
<211> 30
<212> PRT
<213> Homo sapien

<400> 146
Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Gln
1 5 10 15

<213> Homo sapien

<400> 150
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15
Ser Leu Arg Leu Ser Cys Ala Ala Cys Gly Phe Thr Phe Ser
20 25 30

<210> 151
<211> 14
<212> PRT
<213> Homo sapien

<400> 151
Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser
1 5 10

<210> 152
<211> 32
<212> PRT
<213> Homo sapien

<400> 152
Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln
1 5 10 15
Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg
20 25 30

<210> 153
<211> 11
<212> PRT
<213> Homo sapien

<400> 153
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser

<212> PRT
<213> Mus musculus

<400> 157
Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser
1 5 10

<210> 158
<211> 30
<212> PRT
<213> Mus musculus

<400> 158
Gln Val Gln Leu Lys Glu Ser Gly Pro Gly Leu Val Ala Pro Ser Gln
1 5 10 15
Ser Leu Ser Ile Thr Cys Thr Val Ser Gly Phe Ser Leu Thr
20 25 30

<210> 159
<211> 14
<212> PRT
<213> Mus musculus

<400> 159
Trp Val Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Leu Gly
1 5 10

<210> 160
<211> 32
<212> PRT
<213> Mus musculus

<400> 160
Arg Leu Ser Ile Ser Lys Asp Gln Ser Lys Ser Gln Val Phe Leu Lys
1 5 10 15

<213> Mus musculus

<400> 164
Lys Ala Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
1 5 10 15
Leu Ser Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Tyr Cys Ala Arg
20 25 30

<210> 165
<211> 11
<212> PRT
<213> Mus musculus

<400> 165
Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser
1 5 10

<210> 166
<211> 30
<212> PRT
<213> Mus musculus

<400> 166
Gln Val Gln Leu Gln Gln Pro Gly Ala Glu Leu Val Lys Pro Gly Ala
1 5 10 15
Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Tyr Thr Pro Thr
20 25 30

<210> 167
<211> 14
<212> PRT
<213> Mus musculus

<400> 167
Trp Val Lys Gln Arg Pro Gly Gln Gly Leu Glu Trp Ile Gly

<212> PRT
<213> Mus musculus

<400> 171
Trp Val Lys Gln Arg Pro Glu Gln Gly Leu Glu Trp Ile Gly
1 5 10

<210> 172
<211> 32
<212> PRT
<213> Mus musculus

<400> 172
Lys Ala Thr Ile Thr Ala Asp Thr Ser Ser Asn Thr Ala Tyr Leu Gln
1 5 10 15
Leu Ser Ser Leu Thr Ser Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg
20 25 30

<210> 173
<211> 11
<212> PRT
<213> Mus musculus

<400> 173
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
1 5 10

<210> 174
<211> 30
<212> PRT
<213> Mus musculus

<400> 174
Glu Val Lys Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

<213> Mus musculus

<400> 178
Glu Val Lys Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15
Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Phe Asp Phe Ser
20 25 30

<210> 179
<211> 14
<212> PRT
<213> Mus musculus

<400> 179
Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Ile Gly
1 5 10

<210> 180
<211> 32
<212> PRT
<213> Mus musculus

<400> 180
Lys Phe Ile Ile Ser Arg Asp Gln Ala Lys Gln Thr Leu Tyr Leu Gln
1 5 10 15
Met Ser Lys Val Arg Ser Glu Asp Thr Ala Leu Tyr Tyr Cys Ala Arg
20 25 30

<210> 181
<211> 11
<212> PRT
<213> Mus musculus

<400> 181
Trp Gly Gln Gly Thr Thr Val Thr Val Ser Ser

<212> PRT
<213> Mus musculus

<400> 185
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
1 5 10

<210> 186
<211> 30
<212> PRT
<213> Mus musculus

<400> 186
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Lys Pro Gly Gly
1 5 10 15
Ser Leu Lys Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser
20 25 30

<210> 187
<211> 14
<212> PRT
<213> Mus musculus

<400> 187
Trp Val Arg Gln Thr Pro Glu Lys Arg Leu Glu Trp Val Ala
1 5 10

<210> 188
<211> 32
<212> PRT
<213> Mus musculus

<400> 188
Arg Phe Thr Ile Ser Arg Asp Gln Ala Lys Gln Thr Leu Tyr Leu Gln
1 5 10 15

<213> Mus musculus

<400> 192
Lys Thr Thr Leu Thr Val Asp Lys Ser Ser Ser Thr Ala Tyr Met Gln
1 5 10 15
Leu Arg Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala Arg
20 25 30

<210> 193
<211> 11
<212> PRT
<213> Mus musculus

<400> 193
Trp Gly Gln Gly Thr Thr Leu Thr Val Ser Ser
1 5 10

<210> 194
<211> 31
<212> PRT
<213> Mus musculus

<400> 194
Ser Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Lys Pro Gly
1 5 10 15
Gly Ser Val Lys Leu Ser Cys Lys Ala Ser Gly Phe Thr Phe Ser
20 25 30

<210> 195
<211> 14
<212> PRT
<213> Mus musculus

<400> 195
Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Gly

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US01/14349

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : C12P 21/04
US CL : 435/70.1

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 435/70.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Information Disclosure Statement for US Patent Application 09/563,222

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Please See Continuation Sheet

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X — Y	MA, J.K-C. et al. Assembly of monoclonal antibodies with OgG1 and IgA heavy chain domains in transgenic tobacco plants. Eur. J. Immunol. 1994 Vol. 24, pages 131-138. See entire document.	1-4, 12-16, 18, 35-37, 42-47 ——— 1-4, 8-24, 35-52
Y	US 5,840,526 A (CASTERMAN et al) 24 November 1998 (24.11.1998) columns 12-13.	1-4, 8-24, 35-52

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 July 2001 (21.07.2001)	27 AUG 2001
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703)305-3230	Authorized officer Thomas Prasthofer Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet) (July 1998)

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US01/14349

Continuation of Item 4 of the first sheet:
The title is not precise, PCT Rule 4.3, suggested new title follows: "IMMUNOGLOBULIN BINDING PROTEIN ARRAYS IN PLANT CELLS"

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1.

I. Claims 1-4 and 8-24, drawn to a method for preparing an immunoglobulin binding protein array in plant cells.
II. Claims 5-7, drawn to a method for generating plants and progeny containing polynucleotides encoding IgBP components sufficient to form a functional IgBP.
III. Claims 25-34 and 38-52, drawn to a method of preparing a heavy chain binding protein array in eukaryotic cells.
IV. Claims 35-52, drawn to a method for preparing a plant CHBP array.
V. Claims 53-63, drawn to a CHBP array in eukaryotic cells.
VI. Claim 64, drawn to a composition comprising an array of encapsulated CHBPs.
VII. Claim 65 drawn to a method of preparing a heavy chain binding protein array in eukaryotic cells.
VIII. Claim 66, drawn to a CHBP.

The inventions listed as Groups I-VIII do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: .

The inventions appear to share the common special technical feature of heavey chain binding proteins (CHBPs) made in plant cells comprising an amino acid sequence that is at least 75% identical to a constant region "tailpiece" of an alpha chain and "combining sites" comprise regions that are at least 75% identical to 25 consecutive amino acid portions of either light or heavy chains. The special technical feature is shown by MA et al. (1994) Eur. J. Immunol. 24:131-138 not to be novel and to lack an inventive step.

The Ma et al. reference discloses the production of transgenic tobacco plants that express chimeric monoclonal antibodies (abstract). The Plant G1/A and G2/A "components" of expressed monoclonal antibodies include heavy chains, each comprising an amino acid sequence that is at least 75% identical (i.e. 100%) to a constant region "tailpiece" of an alpha chain (figure 1). The combining sites of the expressed antibodies comprise variable regions that are at least 75% identical (i.e. 100%) to 25 consecutive amino acid portions of either light or heavy chains (figure 1). Plants were transformed with a culture of A. tumefaciens containing the cDNA inserts encoding the different "components" of the antibody or "CHBP" as defined in claim 2 (page 132, column 2, section 2.3). The plants were grown on growth medium to form plants and the plants were screened for expression of immunoglobulin (Ig) chains (section 2.3). Plants expressing Ig chains were crossed, seeds were produced and germinated to produce plants that expressed a variety of chimeric antibodies (section 2.3).

Continuation of B. FIELDS SEARCHED Item3:
CAPLUS, MEDLINE, AGRICOLA on STN (CAS Columbus, OH USA) imminoglobulin?, antibod? , array?, librar?, plant?, constant, heavy

3. Human Heavy Chains

Type	FR1	FR2	FR3	FR4
Group I	QVQLVQSGAEVKKPGASVKVSCKASGYTFT	WVRGAPGQGLEWMG	RVTITADTSTSTAYMELSSLRSEDTAVYYCAR	WGQGTLVTVSS
Group II	QVQLQESGPGLVKPSQTLSLTCTVSGGSVS	WIRQPPGKGLEWIG	RVTISVDTSKNQFSLKLSSVTAADTAVYYCAR	WGQGTLVTVSS
Group III	EVQLVESGGGLVQPGGSLRLSCAACGFTFS	WVRQAPGKGLEWVS	RFTISRDNSKNTLYLQMNSLRAEDTAVYYCAR	WGQGTLVTVSS

4. Mouse Heavy Chains

Group IA	EVQLQESGPSLVKPSQTLSLTCSVTGDSIT	WIRQFPGNKLEWMG	RISITRDTSKNQYFLQLNSVTTEDTATYYCAR	WGQGTTVTVSS
Group IB	QVQLKESGPGLVAPSQSLSITCTVSGFSLT	WVRQPPGKGLEWLG	RLSISKDQSKSQVFLKMNSLQTDDTAMYYCAR	WGQGTSVTVSS
Group IIA	EVQLQQSGPELVKPGASVKISCKASGYTFT	WVKQSPGKSLEWIG	KATLTVDKSSSTAYMQLSSLTSEDSAVYYCAR	WGQGTTVTVSS
Group IIB	QVQLQQPGAELVKPGASVKLSCKASGYTPT	WVKQRPGQGLEWIG	KATLTVDKSSSTAYMQLSSLTSEDSAVYYCAR	WVQGTTVTVSS
Group IIC	EVQLQQSGAELVKPGASVKLSCTASGFNIK	WVKQRPEQGLEWIG	KATITADTSSNTAYLQLSSLTSEDTAVYYCAR	WGQGTLVTVSS
Group IIIA	EVKLVESGGGLVQPGGSLRLSCATSGFTFS	WVRQPPGKALEWIA	RFTVSRDTSQSIIYLQMNALRAEDTAIYYCAR	WGAGTTVTVSS
Group IIIB	EVKLLESGGGLVQPGGSLKLSCAASGFDFS	WVRQAPGKGLEWIG	KFIISRDQAKQTLYLQMSKVRSEDTALYYCAR	WGQGTTVTVSS
Group IIIC	EVKLEESGGGLVQPGGSMKLSCVASGFTFS	WVRQSPEKGLEWVA	RFTISRDDSKSSVYLQMNNLRAEDTGIYYCTT	WGQGTLVTVSS
Group IIID	EVQLVESGGGLVKPGGSLKLSCAASGFTFS	WVRQTPEKRLEWVA	RFTISRDQAKQTLYLQMSSLRSEDTAMYYCAR	WGQGTSVTVSS
Group VA	EVQLQQSGAELVRAGSSVKMSCKASGYTFT	WVKQRPGQGLEWIG	KTTLTVDKSSSTAYMQLRSLTSEDSAVYFCAR	WGQGTTLTVSS
Misc.	SEVQLVESGGGLVKPGGSVKLSCKASGFTFS	WVRQAPGKGLEWVG	RFTISRDNSKSTLYLQMSSLRSEDTATYYCAR	WGQGTTVTVSS

FIG. 1B